

```

/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN37E07"
/notes="end : Sp6"

BASE COUNT      179 a      70 c      85 g      448 t      319 others
ORIGIN

Query Match      2.7%; Score 43.8; DB 12; Length 1101;
Best Local Similarity 41.0%; Pred. No. 8.7;
Matches 111; Conservative 13; Mismatches 147; Indels 0; Gaps 0;

QY  1340 aaataatccagggccccaaggacatcggtatataacataaagaagaatcaaat 1399
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  412 WAAAWMAAAAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  1400 gtcccaaggcatagcatataagaacataagcagctacttaagtgacacttg 1459
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  352 AAAAAAAAAAAAAAAAAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAA
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  1460 tattctacataagatgctgcataagcaccataaccacagcggaacttaattcata 1519
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  292 CANTCTTCNCNANCTCCCAATCATCAATTTTATTAATTAATTAATTAATTAATAA 233
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  1520 ttaacatgaatgttccctataatgcatttcataagaaatcctcatgtgacattcaag 1579
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  232 AAAAAAAAAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAATATCTTA 173
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  1580 aggatcatgttgaagaagaacaaaaa 1610
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  172 ANNANANNTATNNANNAAAAAAAAAA 142
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
LOCUS      A2447091      684 bp      DNA      linear      GSS 04-OCT-2000
DEFINITION 1M0243K14R Mouse 10kb plasmid UUGCIM library Mus musculus genomic
clone UUGCIM0243K14 R. DNA sequence.
ACCESSION  A2447091
VERSION     A2447091.1 GI:10598542
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus.
REFERENCE   1 (bases 1 to 684)
AUTHORS     Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C.,
            Islam,H., Longacre,S., Mahmud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
            and Wright,D., Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah
            Genomic Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: dunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0243 row: K column: 14
            Seq primer: CACACAGGAACAGCATGACGC
            Class: plasmid ends
            High quality sequence stop: 684.
            Location/Qualifiers
                1..684
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUGCIM0243K14"
                /clone_lib="Mouse 10kb plasmid UUGCIM library"

```

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
ligated DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (41473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT      198 a      125 c      123 g      228 t
ORIGIN

Query Match      2.7%; Score 43.6; DB 12; Length 684;
Best Local Similarity 55.2%; Pred. No. 9.9;
Matches 85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY  1455 acttgatttccataagatgctgcataagcaccataaccatccagcggaactaatt 1514
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  529 AATTGATTTCATTAATTATTAGTATTAATTAATTAATTAATTAATTAATTAATTAAT
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  1515 tcatttaacattatgttccctataatgcatttcataagaaatcctcatgtgacatt 1574
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  589 TCAAAATTTCAAGTCTCATGCTATCTAATTAATGATGCGATTGGAACCACTGGGACCTTT 648
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY  1575 tcaagagatatgttgaagaagaacaaaaa 1608
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  649 TTATTTTCATTTTTTTAAACATTAATAAACACA 682
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
LOCUS      CNS01090      524 bp      DNA      linear      GSS 12-MAY-2000
DEFINITION Tetradon nigroviridis genome survey sequence PUC-ORI end of clone
196624 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION  AL167541.1 GI:7805598
VERSION     AL167541.1 GI:7805598
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetradon nigroviridis.
ORGANISM    Tetradon nigroviridis.
REFERENCE   1 (bases 1 to 524)
AUTHORS     Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
            Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
            Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetradon nigroviridis
            Unpublished
            2 (bases 1 to 524)
            Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Queller,F.,
            Saurin,W. and Weissenbach,J.
            Human gene number estimate provided by genome wide analysis using
            Tetradon nigroviridis DNA sequence
            Unpublished
            3 (bases 1 to 524)
            Genoscope.

```

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan

COMMENT

- Web : www.genoscope.cns.fr

Determination of this BAC-end sequence was carried out as part of a

COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).

Location/Qualifiers

Source

```

FEATURES
  source
    location/Qualifiers
      1. .1101
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
        /clone_11b="RPC1-98"
        /clone="BACR29K22"
        /note="end : 17"
  434 a 140 c 25 g 377 t 125 others
BASE COUNT
ORIGIN

```

Query Match	2.6%:	Score 41.8:	DB 12,	Length 1101,
Best Local Similarity	44.0%:	Pred. No. 25,		
Matches 77; Conservative	21;	Mismatches 77;	Indels 0;	Gaps 0;

QY 1435 atgcagctacttaaaatgcaacacctgtatttcctacatagatgctgcatagaaagtccata 149
+ : : | | | | | : ||| : | | : : | : |
Db 983 AYACWTTAATTAAAATTTTTAAAAAAATTTTTTTCACAAATTTTWTATAATTTWAMA 924

QY 1495 aaccatccagcgactatttcataataacatcatgtttccttatatgcatcttcacat 155

D5
923 AAAAAAAAAATTTTAAWWAATAAATTTTTAAAAATTTTTTTTT

D6
1555 gaaatctatattgtacattcaagaaqatagtlltaaaaqaacaaaaaanaa 1600

Db 863 WAATTTTTTTTAAAAAATTWAATTTTAAADADADAAAAAAMDAAAAA 809

RESULT 13

AI592135	356 bp	mRNA	linear	EST 15-MAR-2000
LOCUS				
AI592135				
DEFINITION	mm25a1 v1	Soares mouse	3nbms Mus musculus	cdna clone IMAGE:506050

Accession	AF592135	5' mRNA sequence	mouse	mus musculus	CDNA	clone	IMAGE:596006
-----------	----------	------------------	-------	--------------	------	-------	--------------

```

VERSION 1.1
KEYWORDS EST.
SOURCE house mouse.

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
Mammalia; Proboscidea; Rodentia; Muridae; Murinae; Murini; Mus

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus
AUTHORS	Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.
	(bases 1 to 356)

Underwood, K., Steptoe, M., Thelasing, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter, E., Kohn, S., Shin, T., Jackson, V., Cardenas, M., McCann, P.

TITLE	PI	PI INSTITUTION	PI ADDRESS	PI PHONE	PI FAX	PI E-MAIL	PI URL
The WashU-NCI Mouse EST Project 1999	Waterston, R. and Wilson, R.	Washington University	660 South Euclid Avenue St. Louis, MO 63110	314-772-3700	314-772-3701	waterston@wustl.edu	http://www.wustl.edu/~waterston/

UNPUBLISHED (1999)
Journal
Contact: Marra M/Washu-Mouse EST Project 1999
Washington University School of Medicine
COMMENT

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: mougest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the

This read has been verified to hit its original self in the IMAGE Consortium (info@image.jiml.gov) for further information.

correct orientation)
putative full length read
vector to vector length is 357

```
vector to vector length is 327
MGI:363940
```

Seq primer: -40RP from Gibco
POLYA-No.

FEATURES	Location/Qualifiers
source	1. .356

```

/db.xref="taxon:10090"
/clone="IMAGE:598508"
/clone_1lb="Soares mouse 3NbMS"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
lab_host="DH10B"

```

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
TGTTACCAATCATCTGAAGTGGGAGCGCCGCCTGTATTTCCTTTTTTTTTTTTTTTTTT
3'). double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fátima Bonaldo."

BASE COUNT 143 a 43 c 64 g 106 t

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 739)
Genoscope.
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Séquençage ;
BP 191 91006 Evry cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> . This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billard at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC

project grant. The DNA was prepared from embryos by Alain Bucheton and Geneviève Payan. It has been constructed in the vector pBelobAC11.

FEATURES	source	Location/Qualifiers
	1. .739	
		/organism="Drosophila melanogaster"
		/plasmid="pbeloBAC11"
		/db_xref="taxon:7227"
		/clone_11b="DrosBAC"
		/clone="BACN14J12"
		/note="end : T7"
BASE COUNT	130 a 153 c 117 g 229 t . 110 others	
ORIGIN		

BASE COUNT	130	a	153	c	117	g	229	t	.110	others
ORIGIN										
Query Match										
Best Local Similarity	2.6%		Score 41.6:		DB 12:		Length 739:			
Matches	80:		Conservative		37:		Mismatches		99:	
									Indels 0:	
									Gaps	

LOCUS	BH034649	806 bp	DNA	linear	GSS 17-JUL-2001			
DEFINITION	RPCT-24-294M13.TJ RPCT-24 Mus musculus genomic clone RPCT-24-294M13							
ACCESSION	/ DNA sequence.							
VERSION	BH034649							
KEYWORDS	BH034649.1	GI:14807705						
SOURCE	GSS.							
ORGANISM	house mouse.							
	Mus musculus							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
REFERENCE	1 (bases 1 to 806)							
AUTHORS	Zhao,S., Niemman,W., Malek,J., Shatsman,S., Akintet,B., Levins,M.,							
	Tesgaye,G., Geer,K., Krol,M., Shavratsbeyan,A., Gebregorgis,E.,							
	Russell,D., de Jong,P. and Fraser,C.M.							
TITLE	Mouse BAC End sequences from library RPCT-24							
JOURNAL	unpublished (1999)							
COMMENT	Other_GSSs: RPCT-24-294M13.TJ							

```
FEATURES      Location/Qualifiers
source        1. . 806
              /organism="Mus musculus"
```

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-294M13"
/clone_1fb="RPCI-24"
/sex="Male"

```

```
/db_xref="taxon:10090"
```

```
/clone="RPCI-24-294M13"
```

```
/clone_lib="RPCI-24"
/sex="Male"
```

```
/sex="Male"  
/cell_type="
```

```

/cell_type= spleen/brain
/note="Vector: pTARBAC1;

```

RPCI-24 Mouse BAC Library produced by Pieter de Jong.

Library was cloned in the pTABAC1 cloning vector at the

BamHI sites using mboI partially digested male C57BL/6J DNA "

DNA. "

a	148	c	164	q	272	t
---	-----	---	-----	---	-----	---

1

BASE COUNT	222 a	148 c	164 g	272 t
ORIGIN				

ORIGIN

Query Match

2.6%; Score 41.6; DB 12; Length 806;

Best Local Similarity 53.0%; Pred. No. 29;
Matches 89; Conservative 0; Mismatches

Matches	89;	Conservative	0;	Mismatches	79;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	-----	--------	----	------	----

857 ttcataattgattcctgaacaagcattgatatgtgcagttatcatccaatttcgcct 916

Db 605 TTTTAAAAATTAAACATAGACCACACTTGATAGTAGCTGTTTATAATGCAGTTAGAACCCT 546

QY 917 ctgaactagcaagcatattaattacacgaatacttggagttccagatgttccc 976

Db 545 CTGAATATATTTATTAATAAATAAGGCCACTTCTGATACTATTATTTTACT 486

977 gtccatttgaaccgcagatacaaatltacaatatgtactggaata 1024

Db 485 CTTAACTGCAAAACACTACAAACAGTATCAGAAGTGATCACTGTATA 438

Search completed: June 12, 2002, 07:56:52
Job time: 24667 sec

Job time: 24667 sec

Description

1	1602.2	99.5	1630	3	AB015878	AB015878	Carlbden
2	277	17.2	1602	3	AB036714	AB036714	Carlydea
3	48.6	3.0	7218	6	166494	166494	Sequence, 14
4	46.2	2.9	1938	6	AB045319	AB045319	Chiropsal
5	46.2	2.9	56153	6	AX251552	AX251552	Sequence
6	43.8	2.7	15548	6	AX347057	AX347057	Sequence
7	43.2	2.7	9733	6	AX345584	AX345584	Sequence
8	42.8	2.7	1141	6	AX083744	AX083744	Sequence
9	42.8	2.7	1240	3	AF104650	AF104650	Sequence
10	42.8	2.7	7001	6	AX347354	AX347354	Lasloglos
11	42.8	2.7	7001	6	AX349075	AX349075	Sequence
12	42.8	2.7	7302	6	AX345613	AX345613	Sequence
13	42.8	2.7	174019	9	AP001538	AP001538	Sequence
14	42.8	2.7	340000	9	AP001674	AP001674	Sequence
15	42.4	2.6	195721	2	AC027059	AC027059	Homo sapi
16	42.2	2.6	96975	2	AC004995	AC004995	Homo sapi
17	42.2	2.6	349980	6	AX344564	AX344564	Homo sapi
18	41.8	2.6	61143	2	AC099802	AC099802	Homo sapi
19	41.8	2.6	200317	9	HS69813	HS69813	Sequence
20	41.6	2.6	254534	2	AC097326	AC097326	Homo sapi
21	41.6	2.6	140934	2	AC0099431	AC0099431	Human DNA
22	41.6	2.6	169728	2	AC026183	AC026183	Homo sapi
23	41.4	2.6	349980	6	AX344565	AX344565	Sequence
24	41.2	2.6	38342	6	AX251504	AX251504	Sequence
25	41.2	2.6	38342	6	AX344503	AX344503	Sequence
26	41.2	2.6	127197	6	AX196294	AX196294	Sequence
27	40.8	2.5	87487	10	MMU427334	MMU427334	Mus muscu
28	40.6	2.5	5216	6	AX281348	AX281348	Mus muscu
29	40.6	2.5	5216	6	AX345249	AX345249	Sequence
30	40.6	2.5	7113	6	AX345707	AX345707	Sequence
31	40.6	2.5	183272	2	AC094428	AC094428	Sequence
32	40.6	2.5	205472	2	AC005506	AC005506	Rattus no
33	40.6	2.5	256172	2	AC005139	AC005139	Plasmodiu
34	40.6	2.5	310779	2	AC005140	AC005140	Plasmodiu
35	40.4	2.5	7037	6	AX346653	AX346653	Sequence
36	40.4	2.5	7195	6	AX277866	AX277866	Sequence
37	40.4	2.5	7195	6	AX322551	AX322551	Sequence
38	40.4	2.5	158660	2	AC068425	AC068425	Mus muscu
39	40.4	2.5	188826	9	AL359380	AL359380	Human DNA
40	40.2	2.5	3678	10	BC011531	BC011531	Human DNA
41	40.2	2.5	303750	1	AP003133	AP003133	Staphyloc
42	40.2	2.5	346900	1	AP003362	AP003362	Staphyloc
43	40	2.5	47703	3	U23511	U23511	Caenorhabd
44	40	2.5	106076	6	AC008465	AC008465	Homo sapi
45	40	2.5	132771	9	AP004241	AP004241	Homo sapi

ALIGNMENTS

Accession	Length	Type	Version
AB015878	1630 bp	mRNA	linear
Caribdea rastonii crt-1 mRNA for toxin-1, complete cds.			INV 06-SEP-2000

AB015878.1 GI:9988401

Carybdea rastoni cDNA to mRNA

Eukaryota; Metazoa; Cnidaria; Cubozoa; Cubomedusae; Carybdeidae; Carybde

1 (sites)
Nagai H

Nakajima, T.
Novel proteaseaceous toxin from the hen follicle (see 68-1004)

	Doc	Commun	325 (3)	500	500 (3000)
<i>Biorhys</i>					
<i>Biorham</i>					
<i>carybdea</i>					
<i>rastoni</i>					

20422301
2 bases 1 to 16301

Nagai, H. and Takuwa, K.
Direct Submission

Bioorganic Research, Wakayamadai 1-1-1, Shimamoto, Osaka 618-8503
Submitted (25 Nov 1998) Hiroshi Nagai, Sankyo Institute for
Pharmaceutical Research, 1-1-1, Shimamoto, Osaka 618-8503

Japan (E-mail: Hiroshi_Nagai@sunitory.co.jp, Tel: 81-75-962-3743,
Fax: 81-75-962-2115).

FEATURES

Location/Qualifiers
1..1630

/organism="Carybdea rastoni"
/db_xref="taxon/78582"

gene

/gene="crt-1"
28..1380

CDS

/codon_start=1
/product="Toxin-1"
/protein_id="BAB12728.1"

/translation="MILKHPMLFVLAITSAGKGRSDVNSLTKVETALKEASGN
EALFALGKGEIOETKPDVGOATKIIISGSAIDKINSGLDCTTISGCDIYAGIA
TTTGGPVGKIGAVASPVSSILSTFGSSAKNSVAAYIDRLSKIRDEAIOHBAAG
RDFEASAFIDVKKOOSNLDSDSIANAVPVKESFISQLESRIISOGAATSLSD
AKRAVDILLYCOLVYMRRETLVDLAIIYRRGNMEHVAENARVKELEAATLDF
LHKLIPEOALIGAVYHPIASBETSKAILNTYKFEVPPVPRIGNRKLFINSWNTY
SICSEAYGNWMEFGCSNVNRPNIKVSMSKSGSFYEMNSDRKILYTHDQGMGCTL
DEPDGDQGHMRPIPLRHGKYVMSKRWRNEMFYMESSAGYIRSWENNPFGQHWSTI
"

BASE COUNT 517 a 314 c 354 g 445 t
ORIGIN

Query Match 99.5%; Score 1602.2; DB 3; Length 1630;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1604; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Oy 1 gcaacagcactcgtggaagagagaccatgattcgaacatcttcctgctcttatt 60
Db 1 GCACAAGCAGCTTGGTGAAGAGACACCATGATTTCTAACAATCTTCTGGCTTTATT 60
Oy 61 gtccttgcaattactctcgcacaacacatggcaaacgctctgagtgcaattccttact 120
Db 61 GTCCTTGCAATTAATCTCTGCAAAACATGCAAAACGCTGATGTCAATCTTACTTACT 120
Oy 121 aaggtgaagacgtccttaaaagagcttcgtgagcaagagagctctttagaggtt 180
Db 121 AAGGTGAAAGTCTGCTTTAAAGAGGCTTCTGTAGCAAGAGGCTCTCTTGAGGCTTTA 180
Oy 181 gaaggtctaaagagagatccagacaaaacagaccgagcttgagaaagccacaataatc 240
Db 181 GAAGGCTTAAAGAGAGATCCAGACAAAACGACCGAGTTGACAAAGCAACAAAATC 240
Oy 241 ctgtgactcgtcgagatcagcttaagaaaataatctcggagagagcaacaaatcatt 300
Db 241 CTGTGACTCTCGAGATCAGCTAGAAAATTAATTTGAGAGTCAACCAAAATCATTT 300
Oy 301 tctgtgtgctgacacattgtgtcaggaattgcaacaacttttgagagccctgtcgagatg 360
Db 301 TCTGTTGCTTCGACATTTGTCAGAGAAATTCGAAACATTTTGAGAGCCCTGTGCGGAGT 360
Oy 361 ggaatcgagcgttagcttcttcttctcaatcctatcattgttcttactggaactca 420
Db 361 GGAATCGAGCCGTAGCTTCTTTGTTTCTTCAATTCATCATTTGTTACTGGAACCTCA 420
Oy 421 gcaaaagactcagttcgtccgtatattgataagacttlaagcaagatcgagatgagcc 480
Db 421 GCAAAAGACTCAGTGTGCGTTATTGATAGAGCTTTAAGCAAGATCGCGATGAGGCC 480
Oy 481 atccaagacatgacagcaggtgcaccaagagatttgcgtgaatcattcgtcattcag 540
Db 481 ATCCAAGACATGACAGAGGTGCCAAGAGAGATTTCGTAATCATCTGCATTCATTTCAG 540
Oy 541 gtcattgaacagcagtcacacttcaagataagcacttaagatgataattgagcgaatgtt 600
Db 541 GTCAATGAACAGCAGTCCATTCATACAGATAGGACCTTAAGTATCATGTGAGGAATGT 600
Oy 601 cctgttataaatcttgaatttatatcgacagcttgagagcgaatttcccaaggcgca 660
Db 601 CCTGTTATAAATTTAGTAAATTTATCGACAGCTTGGAGAGACAGAAATTTCCAGGCCCA 660

```

```

Db 601 CCTGTTATAAATTTAGTAAATTTATCGACAGCTTGGAGAGACAGAAATTTCCAGGCCCA 660
Oy 661 gcaactaccagctttagcagatgcaaaagagcgttgaactcattcgtcctattgca 720
Db 661 GCAACTACCAGCTTAGGATGCAAAAGAGACCGCTTGAATTCATTCGCTATTGTGCA 720
Oy 721 ctgttagtcaatgagaagaaccttgcgtgcagcttggtcattcctcctcagaagaagaa 780
Db 721 CTGTGATGATGAGAGAAACCTTCTGCTGTCACCTTGGCTATTCCTACAGAAAGAAAT 780
Oy 781 gcagaacacgttgcgaagtgtctgtgaaacgcctaaatagggtaaacaaagagctgct 840
Db 781 GCAGAACACGTCGAGAGGTGTCGAAACCGCTAATAGGGTAAACAAAGAGCTAGCTGT 840
Oy 841 gatcccccaggtttcttcaataatgattcttcaaaaagatgataagtgagtgcttatt 900
Db 841 GATCCCCAGGTTTCTTCAATAATGATTCCTAACAAGCATTTGATGAGCGAGTTAT 900
Oy 901 catccaattctgcctcctgaaactagcaaaacaaatlaaatacacaagaatacttgg 960
Db 901 CATCCAATTTCTGCTCTGGAAGCTAGCAAAACCAATATTAAATTACAGAAATACTTGA 960
Oy 961 gtcacagatgttcccgctcctatttgaaacccgacatacaaatlacaaatagtactg 1020
Db 961 GTTCCAGATGTTCCCGCTCTATTGAAACCGCAGATACAAATTTACAAATAGTTACTGG 1020
Oy 1021 aatacctacagtatatgcaagtgagagcttaccatgggaattacatgctcagagctgtct 1080
Db 1021 AATACCTACAGTATATGATGAGAGCTTACATGGAATTCATGTCAGAGCTGTCT 1080
Oy 1081 aacgttcgaaatccaaatcagggtatccaaatgtctgtgtggttttacaacatggag 1140
Db 1081 AACGTCGGAATCCAAATATCAGGATATCCAAATGTCTGATGGGTTTACACCATATGAG 1140
Oy 1141 aatagcagctcgaggaagtgatatacaacaaagatgacaaagagatggagtgact 1200
Db 1141 AATAGCAGCTCGAGGAAGTGTATATATCCAAAGATGACCAAGATGGGATGGGATCT 1200
Oy 1201 ttgagtagagatccagagtgagcaagagacatagaggttcttctttagagcagtggaag 1260
Db 1201 TTGAGTAGAGATCCAGGTGACCAAGGCCATATGAGTTCAATCTTGTGAGCATGGGAG 1260
Oy 1261 tatalgtlaaagctcctaagagtgagcccaactggttcatgtlatatgaatacagtgccagt 1320
Db 1261 TATATGTAAAGCTCTAAAGAGTGCGCCAACTGTGATGATATGATGATCAAGTGCCAGT 1320
Oy 1321 ggtcattcgtcagcttgagaaataatccagagacctcaagagacattggagataacataa 1380
Db 1321 GGTCAATTCGACGCTGGGAAATAATCCAGAGCTCTCAAGGACATTTGGAGTAAACATA 1380
Oy 1381 ttaagaggaatcaacaatgtcccaaaaggaatagaaataagaacatcaagaatgcag 1440
Db 1381 TTAAGAGGAATCAACAATGTCCCAAGGCAATAGAAATATGAACATCAACGAAATGCAG 1440
Oy 1441 tacttaagatgcacactgtatttctatagagatgtgtcatbaaagtcataaacat 1500
Db 1441 TACTTAAGATGCACACTGTATTTCTACATAGATGTGCTATGAAGTCCATTAACCAT 1500
Oy 1501 ccagcggaactaatlcatatlaaataatgatttcccttataatgacatlttcatgaaatc 1560
Db 1501 CCAGCGGACTAATTCATATTAACATTAATGTTCTTAAATGAACATTTTCATGAAATC 1560
Oy 1561 tctatttgcacattcgaagagatagatgttgcgaagaacaaacaaaaa 1607
Db 1561 TCTATTGTGACATTTCAAGAGGATATGTTGAAAGAAACAAATTCAAA 1607

```

RESULT 2

AB036714 1602 bp mRNA linear INV 06-SEP-2000
LOCUS AB036714
DEFINITION Carybdea alata cat-1 mRNA for toxin-A, complete cds.
ACCESSION AB036714
VERSION AB036714.1 GI:9988403


```

FEATURES             source
source
gene                 gene
CDS                  CDS

REFERENCE            REFERENCE
AUTHORS              AUTHORS
TITLE                TITLE
JOURNAL              JOURNAL

LOCUS               AB045319
DEFINITION          AB045319
ACCESSION           AB045319
VERSION             AB045319.1
KEYWORDS             Chiropsalmus quadrigatus cDNA to mRNA.
SOURCE              Eukaryota; Metazoa; Chldaria; Cubozoa; Cubomedusae; Chiropodidae;
ORGANISM             Chiropsalmus.
                     1 (sites)
REFERENCE            Nagai,H., Takawa-Kuroda,K., Nakao,M., Oshiro,N., Iwanaga,S. and
AUTHORS              Nakajima,T.
TITLE                A Novel Protein Toxin from the Deadly Box Jellyfish (Sea Wasp,
JOURNAL              Habu-kurage) Chiropsalmus quadrigatus
REFERENCE            Biosci. Biotechnol. Biochem. 66, 97-102 (2002)
AUTHORS              2 (bases 1 to 1938)
TITLE                Nagai,H.
JOURNAL              Direct Submission
REFERENCE            Submitted (29-JUN-2000) Hiroshi Nagai, Tokyo University of
AUTHORS              Fisheries, 4-5-7, Komaba, Minato-ku, Tokyo 108-8477, Japan
JOURNAL              (E-mail:nagai@tokyo-u-fish.ac.jp, Tel.:+81-3-5463-0454,
REFERENCE            Fax:+81-3-5463-0398)
FEATURES             location/Qualifiers
source               1..1938
                    /organism="Chiropsalmus quadrigatus"
                    /db_xref="taxon:130731"
                    76..1464
                    /gene="cqt-a"
                    76..1464
                    /gene="cqt-a"
                    /codon_start=1
                    /product="CqTx-A"
                    /protein_id="BAB82520.1"
                    /db_xref="GI:18146993"
                    /translation="MANNYLSLLALFPMGIGASEGTISGLASLAKIDARKPSGK
LEFDYANNOKOIEEKFSNDDERAKVMGIGSLSTANGFOSSGPAINAGCCDLITVG

```

OY	190	aagaagagatccgcagcaacaaaccgagcggattggaacaagcacaaaactcttgatct	249
Dd	241	ATAGAGAAGAAAGTTTCTTAATGACGACCAAGAAGCAAAGTAAGTAATGGCGGTATTGGATCG	300
OY	250	gtcgagtcagctcttagaanaataatcttggagatgtaaccaaatcatcttggttgc	309
Dd	301	TTAAGCACTGCTGTGGGCAGAGTTCCAGTCACAGGTGCCCGCAAAAGATTGTAGTAGATGC	360
OY	310	ctcgacatctgttcgaagatctgcacaaccttggaggccctgtccggatgagtga	369
Dd	361	TTGGATATCTTGTTGGATCTCCTCGCTTTGAAGAAGCTTGCAAAATTTTTTC-----T	414
OY	370	gccgtagctctcttgyttcttccaatctcatcatctgttacttgaagctgcgaagaac	429
Dd	415	CCAATCTTCTCAATCCGTGCAATGTTBTGGGCTCTTTTGGAGACAAGCCGAGAG	474
OY	430	tcaattctgcgcgttatctgatagagccttaaagcaagcatcgagatgagccatccaaaga	489
Dd	475	AGCGTGGGCTCTGTTGTTAAAAAGTAACTCCAGAACATCTGACCACCAAGACTACAGAA	534
OY	490	catgcaagcagctgcgaagagagatttgcatacatctgattcatcatcag---gtcatg	546
Dd	535	GCGCTGTACGGTGTAAAAAGAGATGCAATTTTCCAAAGCATTTTTTGGACGGAGTCGA	594
OY	547	aaacagcagtcacaactctacagatatgagacctaaagtatcatcttgcagcgaaatgtctcgt	606
Dd	595	AATGAACATCTCCGATCTTAGTCCAAACGGAATCTCTCGTTGGGAGCTAAGCTTCTCTGTA	654
OY	607	tataaatctagtaatttatcgcgacgtltgagagcaaat	647
Dd	655	TACCAAGCGCTTCGTTTCATTGCATGGTAACTTCAAAATA	695

RESULT	5			
AX251552/c				
LOCUS		56153 bp	DNA	Linear
DEFINITION	AX251552			PAT 05-OCT-2001
ACCESSION	Sequence 520 from Patent WO0168912.			
VERSION	AX251552			
KEYWORDS	AX251552.1	GI:15984975		
SOURCE				
ORGANISM	synthetic construct.			
	synthetic construct			
	artificial sequence.			
REFERENCE	1 (bases 1 to 56153)			
AUTHORS	Olek,A., Piepenbrock,C. and Berlin,K.			
TITLE	Diagnosis of diseases associated with tumor suppressor genes and oncogenes			
JOURNAL	Patent: WO 0168912-A 520 20-SEP-2001;			
	Epigenomics AG (DE)			
FEATURES	Location/Qualifiers			
source	1..56153			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="chemically treated genomic DNA (Homo sapiens)"			
BASE COUNT	15738 a	1002 c	12377 g	27036 t
ORIGIN				

Query Match 2.7%; Score 46.2; DB 6; Length 56153;
 Best Local Similarity 52.2%; Pred. No. 1.1;
 Matches 128; Conservative 0; Mismatches 113; Indels 4; Gaps 1;

LOCUS AX347057/c 15548 bp DNA linear PAT 01-FEB-2002
 DEFINITION Sequence 2128 from Patent WO0200928.
 ACCESSION AX347057
 VERSION AX347057.1 GI:18494945
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM artificial sequence.
 REFERENCE 1 (sites)
 AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.
 TITLE Diagnosis of diseases associated with the immune system
 JOURNAL Patent: WO 0200928-A 2128 03-JAN-2002;
 Epigenomics AG (DE)
 FEATURES
 source Location/Qualifiers
 1..15548.
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 4209 a 247 c 2903 g 8189 t
 ORIGIN

Query Match 2.7%; Score 43.8; DB 6; Length 15548;
 Best Local Similarity 47.6%; Pred. No. 4.7;
 Matches 129; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

LOCUS AX083744/c 1141 bp DNA linear PAT 28-FEB-2001
 DEFINITION Sequence 22 from Patent WO0111061.
 ACCESSION AX083744
 VERSION AX083744.1 GI:13185472
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM artificial sequence.
 REFERENCE 1 (bases 1 to 1141)
 AUTHORS Kunst, L. and Clemens, S.
 TITLE Regulation of embryonic transcription in plants
 JOURNAL Patent: WO 011061-A 22 15-FEB-2001;
 UNIVERSITY OF BRITISH COLUMBIA (CA)
 FEATURES
 source Location/Qualifiers
 1..1141
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..1141
 /note="consensus sequence of A.t., L.a., and B.n. FAEI promoters"

BASE COUNT 123 a 32 c 42 g 112 t 832 others
 ORIGIN

Query Match 2.7%; Score 42.8; DB 6; Length 1141;

RESULT 7
 LOCUS AX345584/c 9733 bp DNA linear PAT 01-FEB-2002
 DEFINITION Sequence 655 from Patent WO0200928.
 ACCESSION AX345584
 VERSION AX345584.1 GI:18493470
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM artificial sequence.
 REFERENCE 1 (sites)
 AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.
 TITLE Diagnosis of diseases associated with the immune system
 JOURNAL Patent: WO 0200928-A 655 03-JAN-2002;
 Epigenomics AG (DE)
 FEATURES
 source Location/Qualifiers
 1..9733
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 2955 a 93 c 1898 g 4787 t
 ORIGIN

Query Match 2.7%; Score 43.2; DB 6; Length 9733;
 Best Local Similarity 51.7%; Pred. No. 6.7;
 Matches 123; Conservative 0; Mismatches 113; Indels 2; Gaps 1;

LOCUS AX083744/c 1141 bp DNA linear PAT 28-FEB-2001
 DEFINITION Sequence 22 from Patent WO0111061.
 ACCESSION AX083744
 VERSION AX083744.1 GI:13185472
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM artificial sequence.
 REFERENCE 1 (bases 1 to 1141)
 AUTHORS Kunst, L. and Clemens, S.
 TITLE Regulation of embryonic transcription in plants
 JOURNAL Patent: WO 011061-A 22 15-FEB-2001;
 UNIVERSITY OF BRITISH COLUMBIA (CA)
 FEATURES
 source Location/Qualifiers
 1..1141
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 1..1141
 /note="consensus sequence of A.t., L.a., and B.n. FAEI promoters"

BASE COUNT 123 a 32 c 42 g 112 t 832 others
 ORIGIN

Query Match 2.7%; Score 42.8; DB 6; Length 1141;

Db 754 AAAGATAAATCCAGTATCAATGATAGGGGCTTAATGT 713

RESULT 10

AX347354/c 7001 bp DNA linear PAT 01-FEB-2002

LOCUS AX347354 Sequence 1 from Patent WO0200705.

DEFINITION AX347354

ACCESSION AX347354

VERSION AX347354.1

KEYWORDS GI:18495242

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

Source

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 128;

Conservative 0;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Score 42.8;

DB 6;

Length 7001;

Pred. No. 8.6;

Mismatches 142;

Indels 0;

Gaps 0;

Oy 1580 aggatatttgaagaacaaacaaaaaa 1609
 Db 6715 TATATATATTTTACCAATATAACAAA 6686
 RESULT 13
 LOCUS AP001538/c
 DEFINITION Homo sapiens genomic DNA, chromosome 21q21.1-q21.2 clone:B853K11,
 L156-AP region, complete sequence.
 ACCESSION AP001538
 VERSION AP001538.1 GI:7328982
 KEYWORDS HTG.
 SOURCE Homo sapiens DNA, clone:B853K11.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 174019)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.
 TITLE Homo sapiens 174,019 genomic DNA of 21q21.1-q21.2
 JOURNAL Published Only in Database (2000) In press
 REFERENCE 2 (bases 1 to 174019)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Matanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (23-MAR-2000) to the DDBJ/EMBL/GenBank databases.
 Masahira Hattori, The Institute of Physical and Chemical Research
 (RIKEN), Genomic Sciences Center (GSC): Kitasato Univ., 1-15-1
 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
 (E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-42-778-9923, Fax:81-42-778-9924)
 FEATURES
 Source Location/Qualifiers
 1. 174019
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 /clone="B853K11"
 /map="21q21.1-q21.2"
 BASE COUNT 53303 a 31416 c 32248 g 57052 t
 ORIGIN
 Query Match 2.7%. Score 42.8. DB 9: Length 174019;
 Best Local Similarity 47.1%. Pred. No. 6.7;
 Matches 131; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
 Oy 1333 agctgggaataatcccgagcccaaggaattggagatataacataaagaaggaat 1392
 Db 68563 AGAGGAGAGTAAAGTTCACGAGATCTATTGCAACATGATTAATGTTAATGGAAT 68524
 Oy 1393 caacaatgtcccaagaagcatatgaataacatcaacgaatgcagtaactaaatgc 1452
 Db 68523 GATATGTATTAATAAGATTGCTCATAGTCAATGTTCAATGTTCTACAAAATAAA 68464
 Oy 1453 acaactgtattctcaagaatgtcgcatagaaagtcataaacatccagcggaactaa 1512
 Db 68463 AGGATAGTTGTCAGGTACTAATTAATGTAATGCTTACACCATCAATTAATATAA 68404
 Oy 1513 ttcatatcaacattatgttcttataatgaatttcattgaatctctatgttgca 1572
 Db 68403 CATGATTCACAAACATACATATTATGCTATATCTAATCACTATATAAATTTCTTTTCA 68344
 Oy 1573 ttcaagaggatattgttgaagaacaaacaaaaaa 1610
 Db 68343 ATTAAAAAATTAATATTAAAAAGTATACACGAAAA 68306
 RESULT 14
 LOCUS AP001674/c 340000 bp DNA linear PRI 30-MAY-2000
 DEFINITION Homo sapiens genomic DNA, chromosome 21q, section 18/105.
 ACCESSION AP001674 AL163219 BA000005

VERSION AP001674.1 GI:7768666
 KEYWORDS Homo sapiens DNA.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 Hattori,M., Fujiyama,A., Taylor,T.D., Matanabe,H., Yada,T.,
 Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
 Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
 Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
 Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
 Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
 Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
 Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyma,S.,
 Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstieck,G.,
 Hornischler,K., Barand,P., Scharte,M., Schoen,O., Desario,A.,
 Reichelt,J., Kauter,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
 Hennig,S., Rieseemann,L., Dagand,E., Wehmaeyer,S., Borzym,K.,
 Gardiner,K., Mizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
 Vasp,M.L.
 TITLE Direct Submission
 JOURNAL Submitted (10-APR-2000) to the DDBJ/EMBL/GenBank databases. The
 Chromosome 21 Mapping and Sequencing Consortium. * RIKEN Genomic
 Sciences Center, Human Genome Research Group * Institute of
 Molecular Biotechnology, Genome Analysis * Keio University School
 of Medicine, Dept. of Molecular Biology * GSF, Dept. of Genome
 Analysis * Max-Planck Institute for Molecular Genetics (addresses
 see below)
 On May 30, 2000 this sequence version replaced gi:7717271.
 The chromosome 21 mapping and sequencing consortium consisting of
 * RIKEN Genomic Sciences Center, Human Genome Research Group, *
 Sagamihara 228-8555, Japan,
 * e.mail: hattori@gs.c.riken.go.jp/
 * URL: http://hgp.gsc.riken.go.jp/
 * e.mail: hattori@gs.c.riken.go.jp/
 * URL: http://genome.imb-jena.de/
 * Institute of Molecular Biotechnology, Genome Analysis, *
 Beutenbergstrasse 11, D-07745 Jena, Germany,
 * e.mail: gscj-submit@genome.imb-jena.de
 * URL: http://genome.imb-jena.de/
 * Keio University School of Medicine, Molecular Biology, * Tokyo
 160-8582, Japan,
 * e.mail: nshimizu@med.keio.ac.jp
 * URL: http://www.dmb.med.keio.ac.jp/
 * and
 * GSF, Dept. of Genome Analysis,
 * Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
 info.genome@gbf.de
 * URL: http://genome.gbf.de/
 * and
 * Max-Planck Institute for Molecular Genetics,
 * Inestrasse 73, D-14195 Berlin, Germany,
 * e.mail: info-chr21@molgen.mpg.de
 * URL: http://chr21.rz-berlin.mpg.de/
 AL163219: Submitted (10-APR-2000).
 FEATURES
 Source Location/Qualifiers
 1. 340000
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 <1. 6149
 /note="Accession No. AP001537"
 /organism="Homo sapiens"

```
/db_xref="taxon:9606"
/chromosome="21"
/map="21q21.1"
/clone="R749F18, 5' partial"
/clone_lib="RPCT-11 BAC library"
<!.165231
/note="Accession No. AP001538"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q21.1"
/clone="R853K11, 5' partial"
/clone_lib="RPCT-11 BAC library"
15..444
/note="LIM4A"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
555..3531
/note="LIM4A"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
3770..3876
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(5645..6178)
/note="LIM2"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
6181..6534
/note="HEIB"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
8021..8370
/note="HEIB"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(8373..10032)
/note="LIM2"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
10213..10520
/note="AluDb"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
15810..15851
/note="(CA)n"
/rpt_family="Simple_repeat"
/rpt_type= tandem
complement(16628..16697)
/note="MER53"
/rpt_family="DNA"
/rpt_type=DISPERSED
16638..16715
/note="MER53"
/rpt_family="DNA"
/rpt_type=DISPERSED
18847..19585
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
19795..19917
/note="(TA)n"
/rpt_family="Simple_repeat"
20023..20166
/note="(TA)n"
/rpt_family="Simple_repeat"
20289..20448
/note="(TA)n"
/rpt_type= tandem
/rpt_type= tandem
```

```
repeat_region 20467..20542
/note="(TA)n"
/rpt_family="Simple_repeat"
/rpt_type= tandem
complement(21360..22028)
/note="LIM3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(22142..22448)
/note="AluX"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(22519..22891)
/note="MER52A"
/rpt_family="LTR/MER4-group"
/rpt_type=DISPERSED
22976..23239
/note="LIM3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(23244..23519)
/note="MER4C"
/rpt_family="LTR/MER4-group"
/rpt_type=DISPERSED
complement(23644..23861)
/note="MER4C"
/rpt_family="LTR/MER4-group"
/rpt_type=DISPERSED
23925..25091
/note="MER52A"
/rpt_family="LTR/MER4-group"
/rpt_type=DISPERSED
complement(25227..25672)
/note="MER42"
/rpt_family="LTR/MER4-group"
/rpt_type=DISPERSED
complement(25682..25987)
/note="AluY"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
26744..26849
/note="MER46B"
/rpt_family="DNA/MER2-type"
/rpt_type=DISPERSED
complement(26852..27629)
/note="LTR17"
/rpt_family="LTR/Retroviral"
/rpt_type=DISPERSED
complement(27630..27921)
/note="HERV17"
/rpt_family="LTR/Retroviral"

Query Match 2.7%; Score 42.8; DB 9; Length 340000;
Best Local Similarity 47.1%; Pred. No. 6.4;
Matches 131; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

OY 1333 accttgggaataatccaggaacttggagatataacataaagaaggaat 1392
DB 59794 ACAGAGGAGGTAAGTTCACGAGTCTATTGTGCCAACATGATACATTAATGAAT 59735
OY 1393 caacaatgtcccaaggcatcagatatagaacatacgaatgcagtaactaaagtgc 1452
DB 59734 GATATGTGATTTAAAGATGCTCATAGTAGTCGCAATGTTCATTTACAAAATTA 59675
OY 1453 acaacttgattctacataagatgtcgtcatgaagtcataaaccatccagcggaacta 1512
DB 59674 AAGATAGTTGTGACGTACTAATATGCTAAGCTGATTTCCACCATCCATAAATTA 59615
OY 1513 ttcatataaacaataatgttctctataagaactttcaagaacatctctatgtgaca 1572
DB 59614 CATGATCAAAACATACATATTATGATATATCATATCACTAATAAATTCCTTTTCA 59555
OY 1573 ttcaagagatgtttgaagaacaaacaaaaaa 1610
```

```

DB 59554 ATTTAAAAATATATTTAAAAAGTATACCAAGAAA 59517
|||||
RESULT 15
AC027059/C 195721 bp DNA linear HTG 19-APR-2000
LOCUS Homo sapiens chromosome 4 clone RP11-595J23 map 4, WORKING DRAFT
DEFINITION
AC027059
AC027059.2 GI:7596861
VERSION HTG: HTGS-PHASE1, HTGS-DRAFT.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 195721)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 4, clone RP11-595J23
Unpublished
2 (bases 1 to 195721)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dekrelano,K., Dewar,K., Diaz,J.S.,
Dodghe,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,
Gland-pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGuck,A., McKernan,K., McPheters,R.,
Mudrilm,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teefaye,S., Theodore,J., Tyrrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (26-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 19, 2000 this sequence version replaced g1:7329419.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: I7408
Center clone name: 599_J23
-----
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 184383 bases at least Q40
Consensus quality: 191056 bases at least Q30
Consensus quality: 193091 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 194121; sum-of-contigs
Quality coverage: 4.6 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1652: contig of 1652 bp in length
1653 1752: gap of 100 bp
1753 6562: contig of 4810 bp in length
6563 6662: gap of 100 bp
6663 10464: contig of 3802 bp in length
10465 10564: gap of 100 bp
10565 15408: contig of 4844 bp in length
15409 15508: gap of 100 bp
15509 21412: contig of 5904 bp in length
21413 21512: gap of 100 bp
21513 27609: contig of 6097 bp in length
27610 27709: gap of 100 bp
27710 32894: contig of 5185 bp in length
32895 32994: gap of 100 bp
32995 37528: contig of 4534 bp in length
37529 37628: gap of 100 bp
37629 45785: contig of 8157 bp in length
45786 45885: gap of 100 bp
45886 58042: contig of 12157 bp in length
58043 58142: gap of 100 bp
58143 73270: contig of 15127 bp in length
73270 73369: gap of 100 bp
73370 89506: contig of 16137 bp in length
89507 89606: gap of 100 bp
89607 104179: contig of 14573 bp in length
104180 104279: gap of 100 bp
104280 122927: contig of 18648 bp in length
122928 123027: gap of 100 bp
123028 144803: contig of 21776 bp in length
144804 144903: gap of 100 bp
144904 169867: contig of 24964 bp in length
169868 169967: gap of 100 bp
169968 195721: contig of 25754 bp in length.
Location/Qualifiers
1. 195721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-595J23"
/clone_lib="RPC1-11 Human Male BAC"
1. 1652
/note="assembly-fragment"
1753. 6562
/note="assembly-fragment"
6563. 10464
/note="assembly-fragment"
10565. 15408
/note="assembly-fragment"
15509. 21412
/note="assembly-fragment"
21513. 27609
/note="assembly-fragment"
27710. 32894
/note="assembly-fragment"
32995. 37528
/note="assembly-fragment"
37629. 45785
/note="assembly-fragment"
45886. 58042
/note="assembly-fragment"
58143. 73269
/note="assembly-fragment"
73370. 89506
/note="assembly-fragment"
89607. 104179
/note="assembly-fragment"
clone_end:sp6
vector_side:right"

```

```

misc_feature      104280..122927
                  /note="assembly_fragment"
misc_feature      123028..144803
                  /note="assembly_fragment"
misc_feature      144904..169867
                  /note="assembly_fragment"
misc_feature      169968..195721
                  /note="assembly_fragment"
BASE COUNT      61906 a 35617 c 35577 g 61019 t 1602 others
ORIGIN

```

```

Query Match      2.6%; Score 42.4; DB 2; Length 195721;
Best Local Similarity 45.1%; Pred. No. 8.3;
Matches 157; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

OY 845 ccctagatttctcattcaatgatctcctgaacagcatgtatagtgcagtttatcatc 904
    || || || || || || || || || || || || || || || || || || || ||
DB 85073 CCGTATTATTAGTATAGTATGTTAATGTTATGTTATTTATTTACACATTTTGACATA 85014

OY 905 caattctgcctctgaaactgcaaaagaataataattacaagaataacttggagttc 964
    || || || || || || || || || || || || || || || || || || || ||
DB 85013 TATTATGTGTATTAAGCCACAAATAACAGACTACTTAATGAGAGTATGTTTCAACAG 84954

OY 965 cagatgtcccccctcctcttggaaacgcgacatatacaaatltacaatagttacttggaa 1024
    || || || || || || || || || || || || || || || || || || || ||
DB 84953 TAAATGCCATTATCTTCTTAAGTTAAATTAGACTCTATGTTACTAATAGTTAGACAGTA 84894

OY 1025 cctacagatlatgcaatgagccttacatgaggaattacatgttcagagcctgttctaag 1084
    || || || || || || || || || || || || || || || || || || || ||
DB 84893 TTTCAGAGAGGCTAAACTTCTTATTAAGTCAAAATAGCAGTTTCTCTTTGGTGAATTA 84834

OY 1085 ttcggaatccaaatccaggtatccaaatgtctgaltgggttttacaccatggagaa 1144
    || || || || || || || || || || || || || || || || || || || ||
DB 84833 TTTTAAATACAAATTAACATGCTTTCTTAAGTATGAGATGCGGACTTAAAGACTATGCTACT 84774

OY 1145 gcgatcgaggaagtgtatatacccaagcatgaccaaggaatgggat 1192
    || || || || || || || || || || || || || || || || || || || ||
DB 84773 GTACTAAGTGAATTAATTACTGAGGAGAGAGAAAGGAAGGACT 84726

```

Search completed: June 12, 2002, 18:00:40
 Job time: 59091 sec

DR WPI: 1999-580740/49.
 DR P-PSDB: AAY33651.

XX Protein with hemolytic activity, useful for drugs treating jelly fish
 PT stings, pharmaceuticals with blood platelet agglutination activity,
 PT pesticides by use of the hemolytic activity, and study of the hemolytic
 mechanism

PS Disclosure: Page 23-25; 32pp; Japanese.

XX This invention describes a novel protein which has hemolytic activity,
 CC blood platelet agglutination activity and a molecular weight of about
 CC 50,000 Da (by SDS-PAGE). The protein, homologs and analogs of the
 CC invention can be used as a novel approach to developing drugs useful for
 CC the treatment of jelly fish stings, pharmaceuticals with blood platelet
 CC agglutination activity, pesticides by use of the hemolytic activity, and
 CC in the study of the hemolytic mechanism. This sequence encodes the
 CC hemolytic protein described in the invention

XX Sequence 1610 BP; 509 A; 312 C; 352 G; 437 T; 0 other;

Query Match 100.0%; Score 1610; DB 20; Length 1610;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 gcaacagcacttggtgaagagacacacatgctcgaacatctctgctcttatt 60
DB 1 gcaacagcacttggtgaagagacacacatgctcgaacatctctgctcttatt 60
QY 61 gtccttgcaatctctcgcacaaacatgcaacgctcgtatctcaattcttact 120
DB 61 gtccttgcaatctctcgcacaaacatgcaacgctcgtatctcaattcttact 120
QY 121 aagtagaactgctttaaagaagctctctgtagcaacgagcgtcttgaggttta 180
DB 121 aagtagaactgctttaaagaagctctctgtagcaacgagcgtcttgaggttta 180
QY 181 gagggcttaaaagagagatccagacaaacagacgagttgagacaaacaaatc 240
DB 181 gagggcttaaaagagagatccagacaaacagacgagttgagacaaacaaatc 240
QY 241 ctgtagatctgctgagctcctctagaaatctctgtagatgagacccaatctt 300
DB 241 ctgtagatctgctgagctcctctagaaatctctgtagatgagacccaatctt 300
QY 301 tctggttcctcgacatggttcaggaattgcaacaactttgagagccctgctgga 360
DB 301 tctggttcctcgacatggttcaggaattgcaacaactttgagagccctgctgga 360
QY 361 ggaatcggagccgttagctcttcttctcaattctcaatcttcttcttcttct 420
DB 361 ggaatcggagccgttagctcttcttctcaattctcaatcttcttcttcttct 420
QY 421 gcaaggaactcagctgctgcgcttatatagagctttaaagacatcgatgagcc 480
DB 421 gcaaggaactcagctgctgcgcttatatagagctttaaagacatcgatgagcc 480
QY 481 atccaaagacatgacagcaggttcgaagagatttgcctgaatcctgcattctag 540
DB 481 atccaaagacatgacagcaggttcgaagagatttgcctgaatcctgcattctag 540
QY 541 gtcatgaacacagcagtcacatctacagatagcgaactatgacagcgaatgtt 600
DB 541 gtcatgaacacagcagtcacatctacagatagcgaactatgacagcgaatgtt 600
QY 601 cctgtttataaatttagtaatttatcggaacagttgagagcagaattcccaaggcga 660
DB 601 cctgtttataaatttagtaatttatcggaacagttgagagcagaattcccaaggcga 660
QY 661 gcaactacacagcttagcagatgcaagagagcgttgaacttactctgctcatgttaa 720
DB 661 gcaactacacagcttagcagatgcaagagagcgttgaacttactctgctcatgttaa 720

```

```

QY 721 ctgttagcatgagaagaaccttgctgctgactggtgctattctctacagaaagaat 780
DB 721 ctgttagcatgagaagaaccttgctgctgactggtgctattctctacagaaagaat 780
QY 781 gcagaacagctggcaagctgctggtgaacacgttaattggtgtaaacaaagcctagct 840
DB 781 gcagaacagctggcaagctgctggtgaacacgttaattggtgtaaacaaagcctagct 840
QY 841 gatccctaatattctctcgttaattgattctcctggaacacattgagttgcaattat 900
DB 841 gatccctaatattctctcgttaattgattctcctggaacacattgagttgcaattat 900
QY 901 catcaattctgctcctggaactagcaagaacatataattacacgaatactcttga 960
DB 901 catcaattctgctcctggaactagcaagaacatataattacacgaatactcttga 960
QY 961 gttccagatgtcccgctcctatttgaacacgcagatatacaaatltaacaaatgatt 1020
DB 961 gttccagatgtcccgctcctatttgaacacgcagatatacaaatltaacaaatgatt 1020
QY 1021 aataccctacagatatacagctgagcttaccatggaatattcattgcaagctgttct 1080
DB 1021 aataccctacagatatacagctgagcttaccatggaatattcattgcaagctgttct 1080
QY 1081 aacgttcggaatcccaatatacagaggtatccaaaatgctcgtatggtttacacatgag 1140
DB 1081 aacgttcggaatcccaatatacagaggtatccaaaatgctcgtatggtttacacatgag 1140
QY 1141 aatagcagctcggaggaagtgtatatacacaagcatgacaaagatggtggatgct 1200
DB 1141 aatagcagctcggaggaagtgtatatacacaagcatgacaaagatggtggatgct 1200
QY 1201 ttgagatagatccaggttgcacaaagcagatagatgattctcttgaagactggaag 1260
DB 1201 ttgagatagatccaggttgcacaaagcagatagatgattctcttgaagactggaag 1260
QY 1261 tatactgtaagctctaaagagtggtcccaactggttcatgtatatagtataagtgccag 1320
DB 1261 tatactgtaagctctaaagagtggtcccaactggttcatgtatatagtataagtgccag 1320
QY 1321 ggtcaacttcgcagcttggtgaataataatccaggaactcaaggaacttggatataac 1380
DB 1321 ggtcaacttcgcagcttggtgaataataatccaggaactcaaggaacttggatataac 1380
QY 1381 ttaagaggaatcaacaattcccaaggaatcagatataagaatcaaacgaatgag 1440
DB 1381 ttaagaggaatcaacaattcccaaggaatcagatataagaatcaaacgaatgag 1440
QY 1441 tacttaagtgacacatgttatctacatagagatgctgcattgaaagtcataaacat 1500
DB 1441 tacttaagtgacacatgttatctacatagagatgctgcattgaaagtcataaacat 1500
QY 1501 ccaggggactcaattcatatataaactaatgttctctataatgcaatttctgaatc 1560
DB 1501 ccaggggactcaattcatatataaactaatgttctctataatgcaatttctgaatc 1560
QY 1561 tctatttgacatttcaagagatgatttgaagaacaaaataaaaaa 1610
DB 1561 tctatttgacatttcaagagatgatttgaagaacaaaataaaaaa 1610

```

RESULT 2

AAI66621
 ID AAI66621: standard; DNM; 2042 BP.

XX AAI66621:

XX 07-JAN-2002 (first entry)

DE C. alata hemolytic protein encoding DNM.

XX Hemolytic protein; nematocyst; sting; drug; blood platelet;

KW agrochemical; ds.
 XX
 OS Carybdea alata.
 XX
 FH Key Location/Qualifiers
 FT CDS 73..1464
 FT /+tag= a
 XX
 PN WO200170799-A1.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-JP02209.
 XX
 PR 21-MAR-2000; 2000JP-0078967.
 XX
 PA (SUNR) SUNTORY LTD.
 PI Nagai H, Kuroda K, Nakajima T;
 XX
 DR WPI: 2001-611481/70.
 DR P-PSDB: AAG65633.
 XX
 PT New polypeptide for treating stings, and for the development of
 PT pharmaceuticals and agrochemicals, and studying hemolysis, comprises
 XX the Carybdea alata-originated protein.
 PS Examples: Page 23-24; 28pp; Japanese.
 XX
 CC The invention provides a polypeptide from nematocyst of C. alata. The
 CC protein comprises hemolytic activity, and has a molecular weight of about
 CC 50,000 Daltons as determined by SDS-polyacrylamide gel electrophoresis
 CC (SDS PAGE). The protein is useful for treating stings, and also in the
 CC development of drugs with blood platelet aggregating effects, for the
 CC development of agrochemicals with hemolytic effects, and as reagents for
 CC studying hemolytic mechanisms. The present sequence represents the
 CC DNA encoding a C. alata protein having hemolytic activity.
 XX
 SQ Sequence 2042 BP; 616 A; 400 C; 424 G; 602 T; 0 other;

Query Match 17.2%; Score 277; DB 22; Length 2042;
 Best Local Similarity 54.1%; Pred. No. 4.7e-68;
 Matches 634; Conservative 0; Mismatches 525; Indels 12; Gaps 3;

QY 221 ttggacaagccacaataatccttgatcgtcgatcagctcgaagaataattctg 280
 DB 299 ttgctaaagatcatcaactatcgcgtcgatcgtcgatcagctcgaagaataattctg 358
 QY 281 gagatgcaacaaatcattctgtgctcgcagatgctgcaagaattgcaacaatt 340
 DB 359 gagatcccttgatgctgctcaggtgctgcaatcattgcaagaattgcaacaatt 418
 QY 341 ttggaggccctgtcgagtggaatcgagccgtagctcttctgttcttcaattctat 400
 DB 419 ttggaggccctgtcgagtggaatcgagccgtagctcttctgttcttcaattctat 478
 QY 401 catgttactggaagctcagcaagaactcagttgctcgtttatgtagactttaa 460
 DB 479 gctcttctcgtgaataatgtagtggaagtgcaatcaaaagttatgtagcagctttaa 538
 QY 461 gcaagcatcgatgagccatccaaagacatgcaagtgccaaagagaatttgctg 520
 DB 539 aagaataatcgatcaagtgatggaagacatgtaaaagagaagaaagacatttaag 598
 QY 521 aatcatcgtcattcattcgtgcaagaacagcagctcgaatttcaagtagcagactaa 580
 DB 599 ccgtatcacccttcgcaacagtgatcaaaagacagaagaatcagcaggttccattgg 658
 QY 581 gtaacatgcaagaaatgctcgtttaaatttgaatttcaacgagacattggaga 640
 DB 659 attcggtcagaagatgctgtagatgcatcttaccacaatgctgctcgtggaga 718

QY 641 gcaagaattcccaagcgcaagaactaccagcttagcagatgcaagaagccgttgact 700
 DB 719 gcaagaataatcgcgctctgttccaccgataaacaatgaaagcaatccatcaatt 778
 QY 701 tcatctcgtcattgtaactgttagcagatgagaagaacctgtgtgtgacttgacta 760
 DB 779 tcatctcgtcattgtaactgttagcagatgagaagaacctgtgtgtgacttgacta 838
 QY 761 tctctcaaggaagaatgcaagacg-----tgcaagtgctgtggaacgta 814
 DB 839 ttctgtaacagcgctcggtgtgtagatgtagatgtagatgtagatgtagatgtagat 898
 QY 815 atagggtaaacaagaagctagctgtgtagatgtagatgtagatgtagatgtagatgtagat 874
 DB 899 gtagcaaaacaagaagctagctgtgtagatgtagatgtagatgtagatgtagatgtagat 958
 QY 875 aacaagcatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 934
 DB 959 agatctctcgtgtgttccactactacccttagcacaacttagcagcagcttgtaga 1018
 QY 935 tatlaattacacgaatacttggatgtagatgtagatgtagatgtagatgtagatgtagat 991
 DB 1019 ttcttaaacctcaaaatatttggatgtagatgtagatgtagatgtagatgtagatgtagat 1078
 QY 992 gcaagatacaaatltaacaatagtagatgtagatgtagatgtagatgtagatgtagatgtagat 1051
 DB 1079 ttattacagaatgcaaaacagcagtagatgtagatgtagatgtagatgtagatgtagatgtagat 1138
 QY 1052 tgggaattacatgctgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1111
 DB 1139 cgggaatacacaatgctgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1198
 QY 1112 aaatcgcagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1171
 DB 1199 agctggaataatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1255
 QY 1172 agcatgacaagaatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1231
 DB 1256 aacacgccaagaatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1315
 QY 1232 tgaagttacatccttggatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1291
 DB 1316 tcaatttaccctttaaacaatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1375
 QY 1292 ggtcatatataatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1351
 DB 1376 acttggttacaatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1435
 QY 1352 gacctcaaggaatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1382
 DB 1436 atccacagggaagaatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagatgtagat 1466

RESULT 3
 AAS46794/c
 ID AAS46794 standard; DNA; 56153 BP.
 XX
 AC AAS46794;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Tumour suppressor gene derived chemically modified sequence #520.
 XX
 KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
 KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200168912-A2.
 XX
 PD 20-SEP-2001.

[illegible]

Db	36003	CATFA	35999
<hr/>			
RESULT	4		
ID	ABLJ34155/c		
AC	ABLJ34155 standard; DNA: 15548 BP.		
XX			
ABLJ34155;			
DT	26-MAR-2002 (first entry)		
XX			
DE	Human immune system associated gene SEQ ID NO: 2128.		
XX			
KW	Human: immune system disease; cytosine methylation; antiasthmatic;		
KW	antiarteriosclerotic; antianaemic; cytostatic; nootropic;		
KW	neuroprotective; anti-HIV; anticoagulant; ophthalmological;		
KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;		
KW	antiflammatory; cancer; eye disease; arteriosclerosis; anaemia;		
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;		
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;		
KV	gene; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200200928-A2.		
XX			
PD	03-JAN-2002.		
XX			
PF	02-JUL-2001; 2001WO-EP07537.		
XX			
PR	30-JUN-2000; 2000DE-1032529.		
XX			
PR	01-SEP-2000; 2000DE-1043826.		
PA	(EPIG-) EPIGENOMICS AG.		
XX			
PI	Olek A., Piepenbrock C., Berlin K.		
XX			
DR	WPI; 2002-130909/17.		
XX			
PT	Nucleic acid comprising fragment of chemically modified gene, useful		
PT	for diagnosis and treatment of diseases associated with abnormal		
PT	cytosine methylation -		
PS	Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German.		
XX			
CC	The present invention provides a number of human immune system associated		
CC	genes which are modified by the methylation of cytosines. The sequences		
CC	can be used in the diagnosis and treatment of immune system disorders,		
CC	including eye diseases such as retinopathy, neovascular glaucoma and		
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid		
CC	leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,		
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel		
CC	diseases. The present sequence is a gene of the invention.		
SQ	Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 other;		
<hr/>			
Query Match	2.7%	Score 43.8;	DB 24; Length 15548;
Best Local Similarity	47.6%;	Pred. No. 0.19;	
Matches 129; Conservative	0;	Mismatches 142;	Indels 0; Gaps
Dy	1340	aaaataacccaggacctcaaggcactgtgagatatatacataattaagaaggaatcaacaat	1399
Db	11887	AAAATTAAATAATTAATTTAAATAATTAATAAATTAATAAATTAATAAATTAATTT	11828
Oy	1400	gtcccacaaagcatcagaaatalaagacatcaaacgaatgcagtacttaagaagtcacacctg	1459
Db	11827	AAAAATTAAGTATAATTAATAATTAATAATTAATAAATTAATAAATTAATAAATTAATAA	11768
Oy	1460	tatttctacataggatgctgcctcatgaagaagtcataaacaccacccagcgactaatlcata	1519
Db	11767	AATATTAATAAATAATTAATTAATTAATTAATAATTAATAAATTAATAAATTAATAAATTAATAA	11708

[illegible]

```

Db    9314 TTCTACTAATTAATAATTAACAAAAAACCTAATATPAACTATCATPAAAATCTTAAAAATGCC 9255
QY    1492 ataacccaccagcgactaatctcatcattaacattaaigtcttcataatgcatctt 1551
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    9254 -AATAAATTAAATTTAACCAATTTTTATPTTACTTAATATPACAATTAATTAATTT 9197
QY    1552 catgaacatctatgtgcacatctccaaggagatatgtttgagaagaaaaaaa 1609
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    9196 TAAATTAATTCATTTCTATCTATPAAAAAATPAACTACTGTPAAAAAATAATACATAAAA 9139

RESULT        6
ABL32711/c
ID            ABL32711 standard; DNA; 7302 BP.
XX
XX            ABL32711;
XX
DT            26-MAR-2002 (first entry)
XX
DE            Human immune system associated gene SEQ ID NO: 684.
XX
KW            Human; immune system disease; cytosine methylation; antiasthmatic;
KW            antiarteriosclerotic; antianemic; cytostatic; noctropic;
KW            neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
KW            antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW            antifibromatous; cancer; eye disease; arteriosclerosis; anemia;
KW            acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
KW            neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW            gene; ds.
XX
XX            Homo sapiens.
XX
OS            WO200200928-A2.
XX
PN            03-JAN-2002.
PD
PF            02-JUL-2001; 2001WO-EP07537.
PR            30-JUN-2000; 2000DE-1032529.
PR            01-SEP-2000; 2000DE-1043826.
XX
XX            (EPIG-) EPIGENOMICS AG.
XX
XX            Olek A, Piepenbrock C, Berlin K;
XX
XX            WPI: 2002-130909/17.
XX
XX            Nucleic acid comprising fragment of chemically modified gene, useful
XX            PT for diagnosis and treatment of diseases associated with abnormal
XX            PT cytosine methylation -
PS
PS            Claim 1; SEQ ID NO 684; 32bp + Sequence Listing; German.
XX
XX            The present invention provides a number of human immune system associated
XX            CC genes which are modified by the methylation of cytosines. The sequences
XX            CC can be used in the diagnosis and treatment of immune system disorders,
XX            CC including eye diseases such as retinopathy, neovascular glaucoma and
XX            CC macular degeneration, arteriosclerosis, anemia, cancer, acute myeloid
XX            CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX            CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX            CC diseases. The present sequence is a gene of the invention.
XX
SO            Sequence 7302 BP; 1703 A; 147 C; 1814 G; 3638 T; 0 other;

Query Match          2.7%; Score 42.8; DB 24; Length 7302;
Best Local Similarity 47.4%; Pred. No. 0.25;
Matches 128; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

OY    1340 aaataatcacggacctaaagacatgttgagtatacataattaagaaggaatcaacaat 1399
      ||||| |||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB    6955 AAAAAAAAAACCCTPAAAAAAAAACTATCAAATAATCTAATTAATAATAAAAAACCCCTPAAAAAAAA 6896

```


CC resistance. The nucleic acids are used to transform plants, and can
CC produce soybean plants having an rhg1 or an Rhg4 SCN resistant allele.
CC The nucleic acids can be used for investigating rhg1 or Rhg4 haplotypes
CC of soybean plants and for introgressing SCN resistance or partial SCN
CC resistance into soybean plants. They can also be used in plant breeding
CC programmes. The invention also relates to proteins encoded by such
CC nucleic acid molecules, as well as antibodies capable of recognising
CC these proteins. The present sequence is a nucleic acid molecule
CC provided in the specification.
XX
SQ Sequence 127197 BP; 43836 A; 20975 C; 20437 G; 41949 T; 0 other;

Query Match 2.6%; Score 41.2; DB 22; Length 127197;
Best Local Similarity 51.8%; Pred. No. 2.9;
Matches 117; Conservative 0; Mismatches 108; Indels 1; Gaps 1;
QY 1372 ataacataaataaggaatcaacaatgtcccaaggcatacgaatataagacatcaa 1431
DB 95052 ataacataaataaggaatcaacaatgtcccaaggcatacgaatataagacatcaa 95111
QY 1432 cgaatgagacttaagtgacactgttctcacaatgagtgctgcatagaagtc 1491
DB 95112 gtctcaaatcagcactaagaataatcctaataaagtgtggtgtaaatc 95171
QY 1492 ataaccatccagcagcactaattcattataaactaattgttcccttaataagcattt 1551
DB 95172 attca-attaagttaatttttgcaataataatacatttcttaacttaactaattt 95230
QY 1552 catgaatcctctatgtgacattcagaagagatagttgaagaa 1597
DB 95231 atttaactcattcgtgccaattataataataattaaagta 95276

RESULT 9

ABL32347/C
ID ABL32347 standard; DNA; 5216 BP.

AC ABL32347;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 320.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; anti-naemic; cytosatic; neutrotic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antirheumatic; antidiabetic; antipsoriatic;
KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

OS Homo sapiens.

PN WO200200928-A2.

PD 03-JAN-2002.

PE 02-JUL-2001; 2001WO-EP07537.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPICENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI: 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation

XX
PS Claim 1: SEQ ID NO 320; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 5216 BP; 1332 A; 73 C; 1196 G; 2615 T; 0 other;

Query Match 2.5%; Score 40.6; DB 24; Length 5216;
Best Local Similarity 46.9%; Pred. No. 0.89;
Matches 127; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
QY 1340 aaaaataatccagagacctcaagacattgagatatacacaataaaggaatcaacaat 1399
DB 5025 AA 4966
QY 1400 gtcccaaggcatacgaatataagacatcaacgaatgcagtaacttaagtgacacttg 1459
DB 4965 AAACTTAAATT 4906
QY 1460 tattctcatagagatgtcgtcagtcacataacacatccagcagcactaattcata 1519
DB 4905 CAAAAAAAAAGAACACACTGAAATTAACGATTAACAACTTCAACTTAATAATT 4846
QY 1520 ttaacatgaatgttcccttaataatgcatlctcatgaatcctcattgtgacattcaag 1579
DB 4845 TAAATTTTCTAATTAACATTAATTAATTAATTTCTCAAAACCTTAATTCATTAATA 4786
QY 1580 aggatattgttgaagaagaacaaaaa 1610
DB 4785 AATATCTAATTAATTAATAATAATTAATAA 4755

RESULT 10

ABL34459/C
ID ABL34459 standard; DNA; 5216 BP.

AC ABL34459;

DT 26-MAR-2002 (first entry)

DE Human metastasis associated gene SEQ ID NO: 12.

XX Metastasis associated gene; cytosatic; gene therapy; cancer;
KW cytosine methylation; gene; ds.

OS Homo sapiens.

PN WO200177376-A2.

PD 18-OCT-2001.

PE 06-APR-2001; 2001WO-EP03970.

PR 06-APR-2000; 2000DE-1019058.

PR 07-APR-2000; 2000DE-1019173.

PR 30-JUN-2000; 2000DE-1032529.

PA (EPIC-) EPICENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI: 2002-010922/01.

XX New nucleic acid derived from chemically treated metastasis genes,
PT

Dy 131 / cagcgcctacatcttcggaaggccgggaataataatccagaagccaacaagaccattcgagatatcaac l3/b
 ::: ::: ::: ::: ::: ::: ::: ::: ::: |::| |::|
Db 183 rrrrfrnrnfnrnfrnrnfnrnfrnrnfrnrnfrnrnfrnrnfrnrcrgargcrgrugrgcnuratarc 242

Query Match	2.2%	Score 34.8;	DB 2;	Length 4165;
Best Local Similarity	51.3%;	Pred. No. 1.9;		
Matches	81;	Conservative	0;	Mismatches 77;
			Indels	0;
			Gaps	0;

RESULT 8
US-09-188-930-27
; Sequence 27, Application US/09188930A

```

: APPLICANT: Murlson, James Greg
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: TITLE OF INVENTION: and Methods For Their use
: FILE REFERENCE: 11000.1011c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: fastseq for Windows Version 3.0
: SEQ ID NO 27
: LENGTH: 778
: TYPE: DNA
: ORGANISM: Rat
US-09-188-930-27

Query Match      2.18;   Score 34;   DB 3;   Length 778;
Best Local Similarity 54.28;   Pred. No. 1.3;

```


QY	660	agcaac	665
Db	270	TGTAC	275

```

1      RESULT 14
2      US-08-952-089A-26
3      ; Sequence 26, Application US/08952089A
4      ; Patent No. 6165749
5      ;
6      ; GENERAL INFORMATION:
7      ; APPLICANT: SAGAWA, HIROAKI
8      ; APPLICANT: UENO, HARUMI
9      ; APPLICANT: OSHIMA, ATSUSHI
10     ; APPLICANT: KATO, IKUMOSHIN
11     ; TITLE OF INVENTION: PLASMID
12     ; NUMBER OF SEQUENCES: 33
13     ; CORRESPONDENCE ADDRESS:
14     ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
15     ; STREET: PO BOX 747
16     ; CITY: FALLS CHURCH
17     ; STATE: VA
18     ;
19     ; COUNTRY: USA
20     ; ZIP: 22040-0747
21     ;
22     ; COMPUTER READABLE FORM:
23     ; MEDIUM TYPE: Floppy disk
24     ; COMPUTER: IBM PC compatible
25     ; OPERATING SYSTEM: PC-DOS/MS-DOS
26     ; SOFTWARE: Patent Release #1.0, Version #1.30
27     ;
28     ; CURRENT APPLICATION DATA:
29     ; APPLICATION NUMBER: US/08/952,089A
30     ; FILING DATE:
31     ; CLASSIFICATION: 435
32     ; ATTORNEY/AGENT INFORMATION:
33     ; NAME: WEINER, MARC S.
34     ; REGISTRATION NUMBER: 32,181
35     ; REFERENCE/DOCKET NUMBER: 1422-0319P
36     ; TELECOMMUNICATION INFORMATION:
37     ; TELEPHONE: 703-205-8000
38     ; TELEFAX: 703-205-8050
39     ; INFORMATION FOR SEQ ID NO: 26:
40     ; SEQUENCE CHARACTERISTICS:
41     ; LENGTH: 1588 base pairs
42     ; TYPE: nucleic acid
43     ; STRANDEDNESS: double
44     ; TOPOLOGY: linear
45     ;
46     ; MOLECULE TYPE: DNA (genomic)
47     ;
48     ; US-08-952-089A-26

```

```

1      RESULT 15
2      US-08-952-089A-29
3      ; Sequence 29, Application US/08952089A
4      ; Patent No. 6165749
5      ;
6      ; GENERAL INFORMATION:
7      ; APPLICANT: SAGAWA, HIROAKI
8      ; APPLICANT: UENO, HARUMI
9      ; APPLICANT: OSHIMA, ATSUSHI
10     ; APPLICANT: KATO, IKUNOSHIN
11     ; TITLE OF INVENTION: PLASMD
12     ; NUMBER OF SEQUENCE: 33
13     ;
14     ; CORRESPONDENCE ADDRESS:
15     ; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
16     ; STREET: PO BOX 747
17     ; CITY: FALLS CHURCH
18     ; STATE: VA
19     ;
20     ; COUNTRY: USA
21     ; ZIP: 22040-0747
22     ;
23     ; COMPUTER READABLE FORM:
24     ; MEDIUM TYPE: Floppy disk
25     ; COMPUTER: IBM PC compatible
26     ; OPERATING SYSTEM: PC-DOS/MS-DOS
27     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
28     ; CURRENT APPLICATION DATA:
29     ; APPLICATION NUMBER: US/08/952,089A
30     ; FILING DATE:
31     ; CLASSIFICATION: 435
32     ; ATTORNEY/AGENT INFORMATION:
33     ; NAME: WEINER, MARC S.
34     ; REGISTRATION NUMBER: 32,181
35     ; REFERENCE/DOCKET NUMBER: 1422-0319P
36     ; TELECOMMUNICATION INFORMATION:
37     ; TELEPHONE: 703-205-8000
38     ; TELEFAX: 703-205-8050
39     ; INFORMATION FOR SEQ ID NO.: 29:
40     ; SEQUENCE CHARACTERISTICS:
41     ; LENGTH: 4146 base pairs
42     ; TYPE: nucleic acid
43     ; STRANDEDNESS: double
44     ; TOPOLOGY: linear
45     ;
46     ; MOLECULE TYPE: DNA (genomic)
47     ;
48     ; US-08-952-089A-29

```

Search completed: June 12, 2002, 17:56:07
Job time: 55673 sec

Result No.	Score	Query Match	Length	DB	ID	Description
1	49.6	3.1	819	12	CNS03V8B	AL262100 Tetradodon
2	49	3.0	976	12	CNS04E5M	AL286627 Tetradodon
3	46.6	2.9	1101	12	CNS0003G	AL063921 Drosophila
4	45.2	2.8	1101	12	CNS0107S	AL098626 Drosophila
5	44.2	2.7	832	12	CNS0110K	AL100526 Drosophila
6	43.8	2.7	964	12	CNS07087	AL414457 T7 end of
7	43.8	2.7	1101	12	CNS0780R	AL108721 Drosophila
8	43.6	2.7	684	12	A2447091	A2447091 Tm023K14
9	42.6	2.6	524	12	CNS01U9O	AL167541 Tetradodon
10	41.8	2.6	622	9	BB795590	BB795590 BB795590
11	41.8	2.6	1064	12	CNS001JY	AL061101 Drosophila
12	41.8	2.6	1101	12	CNS00E3J	AL066257 Drosophila
13	41.6	2.6	356	9	AI592135	AI592135 mr35ell.y
14	41.6	2.6	739	12	CNS0150P	AL105787 Drosophila
15	41.6	2.6	806	12	BH034649	BH034649 RPCR-2A-2
16	41.4	2.6	1201	12	CNS0162O	AL106218 Drosophila
17	41	2.5	563	9	AW463485	AW463485 Bp230012B

C	18	41	2.5	1204	12	CNS01562	AL106628	Drosophila
C	19	40.8	2.5	928	12	CNS013PT	AL103163	Drosophila
C	20	40.8	2.5	942	12	CNS0186S	AL109318	Drosophila
C	21	40.6	2.5	952	12	CNS000678	AL0656567	Drosophila
C	22	40.6	2.5	477	10	B17866499	B1786649	sa149612
C	23	40.6	2.5	728	12	AO2722964	AO2722664	nbxh00028P
C	24	40.4	2.5	595	12	BH532365	BH532365	BOH00355TF
C	25	40.4	2.5	963	9	AL366565	AL366565	AL565655
C	26	40.4	2.5	977	12	CNS017PG	AL108334	Drosophila
C	27	40.2	2.5	304	10	B1465914	B1465914	1ez2d01.x
C	28	40.2	2.5	382	10	BE849307	BE849307	uw05901.y
C	29	40.2	2.5	386	10	B1152557	B1152597	602918467
C	30	40.2	2.5	485	9	AM121038	AM121038	UI-M-BH2
C	31	40.2	2.5	1101	12	CNS00036	AL064084	Drosophila
C	32	40	2.5	245	10	Z29927	Z29927	HHEA25A
C	33	40	2.5	253	12	AO586879	AO586879	RPCI-11-4
C	34	40	2.5	702	12	A2844194	A2844194	2M0143F07
C	35	40	2.5	702	12	BH090861	BH090861	PM0143F07
C	36	39.8	2.5	417	9	AU062065	AU062065	24-2
C	37	39.8	2.5	467	12	CNS01868	AL1062998	Drosophila
C	38	39.8	2.5	915	10	BG520365	BG520365	ps20c09.y
C	39	39.8	2.5	995	12	CNS010DX	AL098847	Drosophila
C	40	39.6	2.5	654	9	AM335399	AM335399	54669 AGS
C	41	39.4	2.4	549	10	BM404277	BM404277	EST578604
C	42	39.4	2.4	605	9	A1977621	A1977621	EST722215
C	43	39.4	2.4	1001	12	CNS0155H	AL105023	Drosophila
C	44	39.4	2.4	1002	9	AL524873	AL524873	AL524873
C	45	39.2	2.4	362	9	AI852655	AI852655	UI-M-BH0-

RESULT	1	ALIGNMENTS
CNS03V8B		
LOCUS	819 bp	DNA
DEFINITION	CNS03V8B	linear
	Tetradodon nigroviridis genome survey sequence T7 end of clone	GSS 18-MAY-2000
	06H15 of library G from Tetradodon nigroviridis, genomic survey	
	sequence.	
ACCESSION	AL262100	
KEYWORDS	AL262100.1	GI:7983726
WORDS	GSS: genome survey sequence.	
SOURCE	Tetradodon nigroviridis.	
ORGANISM	Tetradodon nigroviridis	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;	
	Tetraodontidae; Tetradodon.	
REFERENCE	1 (bases 1 to 819)	
AUTHORS	Roest-Crolius,H., Jalllon,O., Dasilva,C., Fitzames,C., Fisher,C.,	
	Bonneu,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and	
	Weissenbach,J.	
TITLE	Characterization and repeat analysis of the compact genome of the	
JOURNAL	freshwater pufferfish Tetradodon nigroviridis	
REFERENCE	unpublished	
AUTHORS	2 (bases 1 to 819)	
	Roest-Crolius,H., Jalllon,O., Dasilva,C., Bonneu,L., Fisher,C.,	
	Bernot,A., Fitzames,C., Wincker,P., Broctier,P., Quetier,F.,	
	Saurin,W. and Weissenbach,J.	
TITLE	Human gene number estimate provided by genome wide analysis using	
JOURNAL	Tetradodon nigroviridis DNA sequence	
REFERENCE	unpublished	
AUTHORS	3 (bases 1 to 819)	
TITLE	Genoscope.	
JOURNAL	Direct Submission	
REFERENCE	Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases	
AUTHORS	This sequence is a single read and was generated as part of a large	
TITLE	scale clone-end sequencing project of the Tetradodon nigroviridis	
JOURNAL	genome. For more information, please take a look at	
COMMENT	http://www.genoscope.cns.fr/tetradodon .	
FEATURES	location/Qualifiers	
SOURCE	1. 819	
	/Organism="Tetradodon nigroviridis"	

```

/dB_xref="taxon:99883"
/clone_1b="G"
/note="Genoscope sequence ID : COBG061CD08H.p1-end : T7
BASE COUNT      427 a      13 c      72 g      259 t      48 others
ORIGIN

```

Query Match	3.1%	Score 49.6	DB 12	Length 819
Best Local Similarity	47.3%	Pred. No. 0.39		
Matches 115	Conservative 6	Mismatches 122	Indels 0	Gaps 0

QY	1368	gagttacaataat	taaagaagatca	acaatgltcc	aaagcaca	taagataa	gacat	1427
Db	486	GAATTAATAA	NATATN	GAAGAAAA	AAAAATTA	AAAAAGAA	ARCGAA	AAAGAA
QY	1428	caaacgaatgc	agtaactaa	gtcgacac	ctgtatct	cacataag	atgctgc	1487
Db	546	AAAAAATAA	AAAAATTA	AAAAATTA	AAAAATTA	AAAAATTA	AAAAATTA	
QY	1488	gtccataaac	ctcagcgag	ctaattc	atacat	ttaaacct	taagttl	1547
Db	606	ATATTAATA	ATAATA	TATAAT	TATTAAT	TATTAATA	TTTATTAATA	
QY	1548	tttcacgaac	ctctcat	gtgacac	ttccaga	gatatg	tttgaa	1607
Db	666	TAAAAAATA	ATTATATA	TATATA	AAAAAATA	TATTAATA	TTTAAAAA	
QY	1608	aaa	1610					
Db	726	ATA	728					

RESULT	2
CNS04E5M/C	
LOCUS	
DEFINITION	CNS04E5M 976 bp DNA Linear GSS 21-MAY-2000 Tetraodon nigroviridis genome survey sequence T7 end of clone 103P02 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION	
VERSION	
KEYWORDS	GSS; genome survey sequence. Tetraodon nigroviridis.
SOURCE	Tetraodon nigroviridis. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE	1 (bases 1 to 976) Roest-Crollius,H., Jalllon,O., Dasilva,C., Fizames,C., Fisher,C., Bonneau,L., Billault,A., Quetler,F., Saurin,W., Bernot,A. and Weissenbach,J. Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
AUTHORS	Unpublished
JOURNAL	2 (bases 1 to 976) Roest-Crollius,H., Jalllon,O., Dasilva,C., Bonneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetler,F., Saurin,W. and Weissenbach,J. Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
TITLE	Unpublished
JOURNAL	3 (bases 1 to 976)
REFERENCE	Genoscope. Direct Submission Submitted (12-Apr-2000) to the EMBL/Genbank/DBJ databases This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .
AUTHORS	location/Qualifiers
COMMENT	1..976
FEATURES	/organism="Tetraodon nigroviridis"
Source	

```

/db_xref="taxon:99883"
/clone="103p02"
/clone_11b="G"
/not="genoscope sequence ID : COB6103DH01Lp1-end : 77"
BASE COUNT      243 a      151 c      137 g      327 t      118 others
ORIGIN

```

[illegible]

RESULT 3
 CNS00396/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 FEATURES
 source
 1101 bp DNA linear GSS 03-JUN-1999
 Drosophila melanogaster genome survey sequence TET3 end of BAC #
 BACR08X10 of Rpci-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 AL063921
 AL063921.1 GI:4941778
 GSS.
 fruit fly.
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mammoser in Pletier de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named Rpci-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain y2; cn bw sp, the same strain used for the BDGP's
 p1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
 Location/Qualifiers
 1..1101
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone_lib="Rpci-98"

BASE COUNT	ORIGIN
201 a	503 others
64 c	202 t
131 g	
/clone="BACR08K10"	
/note="end : TET3"	

Query Match	2.9%;	Score 46.6;	DB 12;	Length 1101;
Best Local Similarity	16.2%;	Pred. No. 1.9;		
Matches	90;	Conservative 226;	Mismatches 239;	Indels 0;
				Gaps 0.

[illegible]

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Drosophila melanogaster	1 (bases 1 to 1101)	Genoscope.	Direct Submission		
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			Submitted (23-JUL-1999)	Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
			- Web : www.genoscope.cns.fr		
			Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) -		

<http://www.edgip.ebi.ac.uk> -. This *Drosophila melanogaster* BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pbeloBAC11.

FEATURES	source	location/Qualifiers
	1. 1101	
		/organism="Drosophila melanogaster"
		/plasmid="pBelOBAC11"
		/db_xref="taxon:7227"
		/clone_1lb="DrosBAC"
		/clone="BACN03C18"
		/note="end : Sp6"
BASE COUNT	321 a	145 c 170 g 309 t 156 others
ORIGIN		

Query Match	2.8%	Score 45.2	DB 12	Length 1101
Best Local Similarity	30.9%	Pred. No. 4.1		
Matches 82	Conservative 64	Mismatches 119	Indels 0	Gaps 0

QY 1346 acccaggaacctcaaggacattcgatataacaataattaagaagaaatcaaacagtccca 1405
Db 778 ATAAATGCGCTAATTADTRATGRKKGGCAGATTCKMTTTAAAAGGAATTAATAAMCATWH 837
QY 1406 aaaggcatacgaatatataagacalcaaacgaatgcgacttaagaatgcgaactgtatttc 1465
Db 838 GTTMTATTTWTAADDDTCSTFRMYAYATAMRGKSGSKAKATTTWDVTKAM 897
QY 1466 tacataagatgtctcatgaagtcataacacalcggsgactaatlttcataatcaac 1525
Db 898 KANACGCMGTANNTAANTMATTNATTAATRAAMICACTTWARNNWMAWTUWMA 957
QY 1526 atlaaatgttccctaataaycatlfticaatcctcatgtgacattccaagagata 1585
Db 958 WTTVAATATWTATTNAGTAAFTTWTWTTMAAMMATATATRYMRANHMTTTTTNHHTMY 1017
QY 1586 ttttgtaagaagaacaaaaaaaaa 1610
Db 1018 YMTTWTATAVAWMDTPBAAHNYANA 1042

RESULT	5	
CNS0110K/C		
LOCUS		
DEFINITION	CNS0110K	832 bp DNA linear GSS 26-JUL-1999
ACCESSION	Drosophila melanogaster genome survey sequence T7 end of BAC	
VERSION	BA0N06G12 of DrosBAC library from Drosophila melanogaster (fruit	
KEYWORDS	fly), genomic survey sequence.	
SOURCE	AL100526	
ORGANISM	GI:5612137	
REFERENCE	GSS.	
AUTHORS	Fruit fly.	
TITLE	Drosophila melanogaster	
JOURNAL	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
COMMENT	1 (bases 1 to 832)	
	Genoscope.	
	Direct Submission	
	Submitted (23-JUL-1999)	
	Genoscope - Centre National de Sequencage	
	Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
	- web : www.genoscope.cns.fr)	
	Determination of this BAC-end sequence was carried out as part of a	

FEATURES	Location/Qualifiers
source	1. .832

```

/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_1b="DrosBAC"
/clone="BACN06G12"
/notes="end : 77"
BASE COUNT      147 a      111 c      100 g      423 t      51 others
ORIGIN

```

Query Match	2.7%	Score	44.2	DB	12	Length	832
Best Local Similarity	48.2%	Pred	No. 7.1				
Matches	94	Conservative	9	Mismatches	92	Indels	0
						Gaps	0

[illegible]

RESULT	6
CNS07E8R/c	
LOCUS	CNS07EAR 964 bp DNA linear GSS-08-JUL-2001
DEFINITION	T7 end of clone XBD0A002B08 of library XBD0A from strain CBS 94 of <i>Candida tropicalis</i> , genomic survey sequence.

ACCESSION	AL441457
VERSION	AL441457.1
SOURCE	GI:12224683
KEYWORDS	GSS.
ORGANISM	Candida tropicalis.
	Candida tropicalis

ORGANISM
Canada lupulalis
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
1 (bases 1 to 964)

REFERENCE

TITLE	Dujon, B.
	Genomic exploration of the hemiascomycetous yeasts: 16. <i>Candida tropicalis</i>
ISSN	087-41
PAGE	91-94 (2000)
JOURNAL	

MEDLINE REFERENCE AUTHORS
20584726
2 (bases 1 to 964)
Soucieur, J. L., Aigue, M., Artiguenave, F., Blandin, G., Bollotin-Fukuhara, M., Bon, E., Brottier, P., Casaricola, S.,

Malpercu, A., Neveglise, C., Ozler-Kalogeropoulos, O., Potter, S., Saurin, W., Tekala, F., Toffiano-Nioche, C., Wesolowski-Louvel, M., Winckel, P., and Weissendbach, J.

JOURNAL
MEDLINE
REFERENCE
3 (base 1 to 964)
Yeast species for molecular evolution studies
FEBS Lett. 487 (1), 3-12 (2000)
20584711

AUTHORS Genoscope.
TITLE Direct Submisslon
JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail

COMMENT

This GSS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *varium*, *Saccharomyces exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*

FEATURES
source Location/Qualifiers
1..964

```

/organism="Candida tropicalis"
/strain="CBS 94"
/db_xref="taxon:5482"
/clone="XBD0A002B08"
/clone_id="XBD0A"
/note:end : T7"
269 a 106 c 78 g 425 t 86 others
BASE COUNT
ORIGIN

```

Query Match	2.7%	Score 43.8:	DB 12:	length 964:
Best Local Similarity	34.3%	Pred. No. 8:	7:	
Matches	93:	Conservative	48:	Mismatches 130:
				Indels 0:
				Gaps 0:

DQ 1340 aaataatccaggaacctcaagacatctggagtatacatataatlaaaaggatcacaacat 1399
 :
 Db 453 RAATAAATAAAWTTGTTGTGAARAWWGAAAAAAAATWAAAAAMMWRARAARRAAAA 394

393. AWTGAAGAAAACCCWAAAAWAWAWTTTAAAPWTTAAAPARAWTTAAATTAAATTTGT 334

```
Db      : ::| |::| |||||::| |::| ::| ::|
333 CTRNAAWWTTCGWAWTGTATWTTGTAACAAAAAARAAWRWTCGWAAAAMWTATTMAWMW 274
```

[illegible]

Db 213 WWWAAWTTAGCAATATAAAAAAAAAA 183

CNS01807/c				
LOCUS	1101 bp	DNA	Linear	GSS 26-JUL-1999
DEFINITION	CNS01807			
	Proscophila melanoaster genome survey sequence SP6 end of BAC			
	Proscophila melanocaster			
	BAC01807 of Proscophila library from Proscophila melanocaster cfunit			

ACCESSION
AL108721
VERSION
AL108721.1
KEYWORDS
GSS.
11y), genomic survey sequence.

ORGANISM melanogaster
Drosophila
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

REFERENCE
AUTHORS
TITLE
Journal
Submitted (23-III-1999) Genoscope - Centre National de Séquençage :
1 (bases 1 to 1101)
Genoscope.
Direct Submission

COMMENT
 E-mail: genoscope@cns.fr
 Web: www.genoscope.cns.fr
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -

library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton

FEATURES	Location/Qualifiers
source	1..1101
	/organism="Drosophila melanogaster"
	/plasmid="pbeloBac11"